REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* APRIL 2013	**2. REPORT TYPE** JOURNAL ARTICLE (Pre Print)	**3. DATES COVERED** *(From - To)* JAN 2012 – APR 2013

4. TITLE AND SUBTITLE

BLIND ONE-BIT COMPRESSIVE SAMPLING

5a. CONTRACT NUMBER
IN-HOUSE

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
62788F

6. AUTHOR(S)

Lixin Shen (Syracuse University) and Bruce Suter (AFRL)

5d. PROJECT NUMBER
T2AW

5e. TASK NUMBER
IN

5f. WORK UNIT NUMBER
HO

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

PRIME
Air Force Research Laboratory/
Information Directorate
Rome Research Site/RITB
525 Brooks Road
Rome NY 13441-4505

SUB
Syracuse University
900 S Crouse Ave
Syracuse, NY 13210

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/Information Directorate
Rome Research Site/RITB
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RI-RS-TP-2013-036

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA Case Number: 88ABW-2013-0384
DATE CLEARED: 28 JAN 2013

13. SUPPLEMENTARY NOTES
One or more of the authors is a U.S. Government employee working within the scope of their Government job; therefore, the U.S. Government is joint owner of the work and has the right to copy, distribute, and use the work. All other rights are reserved by the copyright owner.

14. ABSTRACT
The problem of 1-bit compressive sampling is addressed in this paper. We introduce an optimization model for reconstruction of sparse signals from 1-bit measurements. The model targets a solution that has the least ℓ0-norm among all signals satisfying consistency constraints stemming from the 1-bit measurements. An algorithm for solving the model is developed. Convergence analysis of the algorithm is presented. Our approach is to obtain a sequence of optimization problems by successively approximating the ℓ0-norm and to solve resulting problems by exploiting the proximity operator. We examine the performance of our proposed algorithm and compare it with the binary iterative hard thresholding (BIHT) [11] a state-of-the-art algorithm for 1-bit compressive sampling reconstruction. Unlike the BIHT, our model and algorithm does not require prior knowledge on the sparsity of the signal. This makes our proposed work a promising practical approach for signal acquisition.

15. SUBJECT TERMS
Compressive Sampling, L0 Norm, Sparse Signals

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** MICHAEL LITTLE
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	25	**19b. TELEPHONE NUMBER** (*Include area code)* N/A

Blind One-Bit Compressive Sampling *

Lixin Shen[†] Bruce W. Suter[‡]

January 17, 2013

Abstract

The problem of 1-bit compressive sampling is addressed in this paper. We introduce an optimization model for reconstruction of sparse signals from 1-bit measurements. The model targets a solution that has the least ℓ_0-norm among all signals satisfying consistency constraints stemming from the 1-bit measurements. An algorithm for solving the model is developed. Convergence analysis of the algorithm is presented. Our approach is to obtain a sequence of optimization problems by successively approximating the ℓ_0-norm and to solve resulting problems by exploiting the proximity operator. We examine the performance of our proposed algorithm and compare it with the binary iterative hard thresholding (BIHT) [11] a state-of-the-art algorithm for 1-bit compressive sampling reconstruction. Unlike the BIHT, our model and algorithm does not require prior knowledge on the sparsity of the signal. This makes our proposed work a promising practical approach for signal acquisition.

1 Introduction

Compressive sampling is a recent advance in signal acquisition [4, 5]. It provides a method to reconstruct a sparse signal $x \in \mathbb{R}^n$ from linear measurements

$$y = \Phi x, \tag{1}$$

where Φ is a given $m \times n$ measurement matrix with $m < n$ and $y \in \mathbb{R}^m$ is the measurement vector acquired. The objective of compressive sampling is to deliver an approximation to x from y and Φ. It has been demonstrated that the sparse signal x can be recovered exactly from y if Φ has Gaussian i.i.d. entries and satisfies the restricted isometry property [5]. Moreover, this sparse signal can be identified as a vector that has the smallest ℓ_0-norm among all vectors yielding the same measurement vector y under the measurement matrix Φ.

However, the success of the reconstruction of this sparse signal is based on the assumption that the measurements have infinite bit precision. In realistic settings, the measurements are never exact and must be discretized prior to further signal analysis. In practice, these measurements are quantized, a mapping from a continuous real value to a discrete value over some finite range. As usual, quantization inevitably introduces errors in measurements. The problem of estimating a sparse signal from a set of quantized measurements has been addressed in recent literature. Surprisedly, it has been demonstrated theoretically and numerically that 1-bit per measurement is enough to retain information for sparse signal reconstruction. As pointed out in [3, 11], quantization

[*] This research is supported in part by an award from National Research Council via the Air Force Office of Scientific Research and by the US National Science Foundation under grant DMS-1115523.

[†] Department of Mathematics, Syracuse University, Syracuse, NY 13244, USA. Email: lshen03@syr.edu.

[‡] Air Force Research Laboratory, AFRL/RITB, Rome, NY 13441-4505. Email: Bruce.Suter@rl.af.mil

to 1-bit measurements is appealing in practical applications. First, 1-bit quantizers are extremely inexpensive hardware devices that test values above or below zeros, enabling simple, efficient, and fast quantization. Second, 1-bit quantizers are robust to a number of non-linear distortions applied to measurements. Third, 1-bit quantizers do not suffer from dynamic range issues. Due to these attractive properties of 1-bit quantizers, in this paper we will develop efficient algorithms for reconstruction of sparse signals from 1-bit measurements.

The 1-bit compressive sampling framework originally introduced in [3] is briefly described as follows. Formally, it can be written as

$$y = A(x) := \operatorname{sign}(\Phi x), \tag{2}$$

where the function $\operatorname{sign}(\cdot)$ denotes the sign of the variable, element-wise, and zero values are assigned to be $+1$. Thus, the measurement operator A, called a 1-bit scalar quantizer, is a mapping from $\mathbb{R}^n$ to the Boolean cube $\mathcal{B}^m := \{-1, 1\}^m$. Note that the scale of the signal has been lost during the quantization process. We search for a sparse signal $x^\star$ in the unit ball of $\mathbb{R}^m$ such that the sparse signal $x^\star$ is consistent with our knowledge about the signal and measurement process, i.e., $A(x^\star) = A(x)$.

The problem of reconstructing a sparse signal from its 1-bit measurements is generally non-convex, and therefore it is a challenge to develop an algorithm that can find a desired solution. Nevertheless, since this problem was introduced in [3] in 2008, there are several algorithms that have been developed for attacking it [3, 13, 19, 21]. Among those existing 1-bit compressive sampling algorithms, the binary iterative hard thresholding (BIHT) [11] exhibits its superior performance in both reconstruction error and as well as consistency via numerical simulations over the algorithms in [3, 13]. When there are a lot of sign flips in the measurements, a method based on adaptive outlier pursuit for 1-bit compressive sampling was proposed in [21]. The algorithms in [11, 21] require the sparsity of the desired signal to be given in advance. This requirement, however, is hardly satisfied in practice. By keeping only the sign of the measurements, the magnitude of the signal is lost. The models associated with the aforementioned algorithms seek sparse vectors x satisfying consistency constraints (2) in the unit sphere. As a result, these models are essentially non-convex and non-smooth. In [19], a convex minimization problem is formulated for reconstruction of sparse signals from 1-bit measurements and is solved by linear programming. The details of the above algorithms will be briefly reviewed in the next section.

In this paper, we introduce a new ℓ_0 minimization model over a convex set determined by consistency constraints for 1-bit compressive sampling recovery and develop an algorithm for solving the proposed model. Our model does not require prior knowledge on the sparsity of the signal, therefore, is referred to as the blind 1-bit compressive sampling model. Our approach for dealing with our proposed model is to obtain a sequence of optimization problems by successively approximating the ℓ_0-norm and to solve resulting problems by exploiting the proximity operator [17]. Convergence analysis of our algorithm is presented.

This paper is organized as follows. In Section 2 we review and comment current 1-bit compressive sampling models and then introduce our own model by assimilating advantages of existing models. Heuristics for solving the proposed model are discussed in Section 3. Convergence analysis of the algorithm for the model is studied in Section 4. A numerical implementable algorithm for the model is presented in Section 5. The performance of our algorithm is demonstrated and compared with the BIHT in Section 6. We present our conclusion in Section 7.

2 Models for One-Bit Compressive Sampling

In this section, we begin with reviewing existing models for reconstruction of sparse signals from 1-bit measurements. After analyzing these models, we propose our own model that assimilates the advantages of the existing ones.

Using matrix notation, the 1-bit measurements in (2) can be equivalently expressed as

$$Y\Phi x \geq 0, \tag{3}$$

where

$$Y := \operatorname{diag}(y)$$

is an $m \times m$ diagonal matrix whose ith diagonal element is the ith entry of y. The expression $Y\Phi x \geq 0$ in (3) means that all entries of the vector $Y\Phi x$ are no less than 0. Hence, we can treat the 1-bit measurements as sign constraints that should be enforced in the construction of the signal x of interest. In what follows, equation (3) is referred to as sign constraint or consistency condition, interchangeably.

The optimization model for reconstruction of a sparse signal from 1-bit measurements in [3] is

$$\min \|x\|_1 \quad \text{s.t.} \quad Y\Phi x \geq 0 \quad \text{and} \quad \|x\|_2 = 1, \tag{4}$$

where $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the ℓ_1-norm and ℓ_2-norm of a vector, respectively. In model (4), the ℓ_1-norm objective function is used to favor sparse solutions, the sign constraint $Y\Phi x \geq 0$ is used to impose the consistency between the 1-bit measurements and the solution, the constraint $\|x\|_2 = 1$ ensures a nontrivial solution lying on the unit ℓ_2 sphere.

Instead of solving model (4) directly, a relaxed version of model (4)

$$\min \left\{ \lambda \|x\|_1 + \sum_{i=1}^{m} h((Y\Phi x)_i) \right\} \quad \text{s.t.} \quad \|x\|_2 = 1 \tag{5}$$

was proposed in [3] and solved by employing a variation of the fixed point continuation algorithm in [10]. Here λ is a regularization parameter and h is chosen to be the one-sided ℓ_1 (or ℓ_2) function, defined at $z \in \mathbb{R}$ as follows

$$h(z) := \begin{cases} |z| \ \ (\text{or } \frac{1}{2}z^2), & \text{if } z < 0; \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

We remark that the one-sided ℓ_2 function was adopted in [3] due to its convexity and smoothness properties that are required by a fixed point continuation algorithm.

In [13] a restricted-step-shrinkage algorithm was proposed for solving model (4). This algorithm is similar in sprit to trust-region methods for nonconvex optimization on the unit sphere and has a provable convergence guarantees.

Binary iterative hard thresholding (BIHT) algorithms were recently introduced for reconstruction of sparse signals from 1-bit measurements in [11]. The BIHT algorithms are developed for solving the following constrained optimization model

$$\min \sum_{i=1}^{m} h((Y\Phi x)_i) \quad \text{s.t.} \quad \|x\|_0 \leq s \quad \text{and} \quad \|x\|_2 = 1, \tag{7}$$

where h is defined by equation (6), s is a positive integer, and the ℓ_0-norm $\|x\|_0$ counts the number of non-zero entries in x. Minimizing the objective function of model (7) enforces the consistency condition (3). The BIHT algorithms for model (7) are a simple modification of the iterative thresholding algorithm proposed in [2]. It was shown numerically that the BIHT algorithms perform significantly better than the other aforementioned algorithms in [3, 13] in terms of both reconstruction error as well as consistency. Numerical experiments in [11] further show that the BIHT algorithm with h being the one-sided ℓ_1 function performs better in low noise scenarios while the BIHT algorithm with h being the one-sided ℓ_2 function perform better in high noise scenarios. Recently, a robust method for recovering signals from 1-bit measurements using adaptive outlier pursuit was proposed for the measurements having noise (i.e., sign flips) in [21].

The algorithms reviewed above for 1-bit compressive sampling are developed for optimization problems having convex objective functions and non-convex constraints. In [19] a convex optimization program for reconstruction of sparse signals from 1-bit measurements was introduced as follows:

$$\min \|x\|_1 \quad \text{s.t.} \quad Y\Phi x \geq 0 \quad \text{and} \quad \|\Phi x\|_1 = p, \tag{8}$$

where p is any fixed positive number. The first constraint $Y\Phi x \geq 0$ requires that a solution to model (8) should be consistent with the 1-bit measurements. If a vector x satisfies the first constraint, so is ax for all $0 < a < 1$. Hence, an algorithm for minimizing the ℓ_1-norm by only requiring consistency with the measurements will yield the solution x being zero. The second constraint $\|\Phi x\|_1 = p$ is then used to prevent model (8) from returning a zero solution, thus, resolves the amplitude ambiguity. By taking the first constraint into consideration, we know that $\|\Phi x\|_1 = \langle y, \Phi x\rangle$, therefore, the second constraint is $\langle \Phi^\top y, x\rangle = p$ which is linear. This confirms that both objective function and constraints of model (8) are convex. It was further pointed out in [19] that model (8) can be cast as a linear program. As comparing model (8) with model (4), both the constraint $\|x\|_2 = 1$ in model (4) and the constraint $\|\Phi x\|_1 = p$ in model (8), the only difference between both models, enforce a non-trivial solution. However, as we have already seen, model (8) with the constraint $\|\Phi x\|_1 = p$ can be solved by a computationally tractable algorithm.

Let us further comment on models (7) and (8). First, the sparsity constraint in model (7) is impractical since the sparsity of the underlying signal is unknown in general. Therefore, instead of imposing this sparse constraint, we consider to minimize an optimization model having the ℓ_0-norm as its objective function. Second, although model (8) can be tackled by efficient linear programming solvers and the solution of model (8) preserves the effective sparsity of the underlying signal (see [19]), the solution is not necessarily sparse in general as shown in our numerical experiments (see Section 6). Motivated by the aforementioned models and the associated algorithms, we plan in this paper to reconstruct sparse signals from 1-bit measurements via solving the following constrained optimization model

$$\min \|x\|_0 \quad \text{s.t.} \quad Y\Phi x \geq 0 \quad \text{and} \quad \|\Phi x\|_1 = p, \tag{9}$$

where p is again a arbitrary positive number. This model has the ℓ_0-norm as its objective function and inequality $Y\Phi x \geq 0$ and equality $\|\Phi x\|_1 = p$ as its convex constraints.

We remark that the actual value of p is not important as long as it is positive. More precisely, suppose that $\mathcal{S}$ and $\mathcal{S}^\diamond$ are two sets collecting all solutions of model (9) with $p = 1$ and $p = p^\diamond > 0$, respectively. If $x \in \mathcal{S}$, that is, $Y\Phi x \geq 0$ and $\|\Phi x\|_1 = 1$, then, by denoting $x^\diamond := p^\diamond x$, it can be verified that $\|x^\diamond\|_0 = \|x\|_0$, $Y\Phi x^\diamond \geq 0$, and $\|\Phi x^\diamond\|_1 = p^\diamond$. That indicates $x^\diamond \in \mathcal{S}^\diamond$. Therefore, we have that $p^\diamond \mathcal{S} \subset \mathcal{S}^\diamond$. Conversely, we can show that $\mathcal{S}^\diamond \subset p^\diamond \mathcal{S}$ by reverting above steps. Hence, $p^\diamond \mathcal{S} = \mathcal{S}^\diamond$. Without loss of generality, the positive number p is always assumed to be 1 in the rest part of the paper.

To close this section, we compare model (7) and our proposed model (9) in the following result.

Proposition 2.1 *Let $y \in \mathbb{R}^m$ be the 1-bit measurements from an $m \times n$ measurement matrix Φ via equation* (2) *and let s be a positive integer. Assume that the vector $x \in \mathbb{R}^n$ is a solution to* *model* (9)*. Then model* (7) *has the unit vector* $\frac{x}{\|x\|_2}$ *as its solution if* $\|x\|_0 \leq s$*; otherwise, model* (7) *can not have a solution satisfying the consistency constraint if* $\|x\|_0 > s$*.*

Proof: Since the vector x is a solution to model (9), then x satisfies the consistency constraint $Y\Phi x \geq 0$. Hence, it, together with definition of h in (6), implies that

$$\sum_{i=1}^{m} h\left(\left(Y\Phi\frac{x}{\|x\|_2}\right)_i\right) = 0.$$

We further note that $\left\|\frac{x}{\|x\|_2}\right\|_0 = \|x\|_0$ and $\left\|\frac{x}{\|x\|_2}\right\|_2 = 1$. Hence, the vector $\frac{x}{\|x\|_2}$ is a solution of model (7) if $\|x\|_0 \leq s$.

On the other hand, if $\|x\|_0 > s$ then all solutions to model (7) do not satisfy the consistency constraint. Suppose this statement is *false*. That is, there exists a solution of model (7), say $x^\sharp$, such that $Y\Phi x^\sharp \geq 0$, $\|x^\sharp\|_0 \leq s$, and $\|x^\sharp\|_2 = 1$ hold. Set $x^\diamond := \frac{x^\sharp}{\|\Phi x^\sharp\|_1}$. Then $\|x^\diamond\|_0 = \|x^\sharp\|_0 \leq s$, $Y\Phi x^\diamond \geq 0$, and $\|\Phi x^\diamond\|_1 = 1$. Since $\|x^\diamond\|_0 < \|x\|_0$, it turns out that x is not a solution of model (9). This contracts our assumption on the vector x. This completes the proof of the result. □

From Proposition 2.1, we can see that the sparsity s for model (7) is critical. If s is set too large, a solution to model (7) may not be the sparsest solution satisfying the consistency constraint; if s is set too small, solutions to model (7) cannot satisfy the consistency constraint. In contrast, our model (9) does not require the sparsity constraint used in model (7) and delivers the sparsest solution satisfying the consistency constraint. Therefore, these properties make our model more attractive for 1-bit compressive sampling than the BIHT. Since sparsity of the underlying signal is not specified in advance in model (9), we refer it to as *blind 1-bit compressive sampling model.*

3 An Algorithm for the Blind 1-Bit Compressive Sampling

In this section, we will develop algorithms for the proposed model (9). We first reformulate model (9) as an unconstrained optimization problem via the indicator function of a closed convex set in $\mathbb{R}^{m+1}$. It turns out that the objective function of this unconstrained optimization problem is the sum of the ℓ_0-norm and the indicator function composing with a matrix associated with the 1-bit measurements. Instead of directly solving the unconstrained optimization problem we use some smooth concave functions to approximate the ℓ_0-norm and then linearize the concave functions. The resulting model can be viewed as an optimization problem of minimizing a weighted ℓ_1-norm over the closed convex set. The solution of this resulting model is served as a new point at which the concave functions will be linearized. This process is repeatedly performed until a certain stopping criteria is met. Several concrete examples for approximating the ℓ_0-norm are provided at the end of this section.

We begin with introducing our notation and recalling some background from convex analysis. For the d-dimensional Euclidean space $\mathbb{R}^d$, the class of all lower semicontinuous convex functions $f : \mathbb{R}^d \to (-\infty, +\infty]$ such that $\mathrm{dom} f := \{x \in \mathbb{R}^d : f(x) < +\infty\} \neq \emptyset$ is denoted by $\Gamma_0(\mathbb{R}^d)$. The indicator function of a closed convex set C in $\mathbb{R}^d$ is defined, at $u \in \mathbb{R}^d$, as

$$\iota_C(u) := \begin{cases} 0, & \text{if } u \in C; \\ +\infty, & \text{otherwise.} \end{cases}$$

Clearly, ι_C is in $\Gamma_0(\mathbb{R}^d)$ for any closed nonempty convex set C.

Next, we reformulate model (9) as an unconstrained optimization problem. To this end, from the $m \times n$ matrix Φ and the m-dimensional vector y in equation (2), we define an $(m+1) \times n$ matrix

$$B := \begin{bmatrix} \operatorname{diag}(y) \\ y^\top \end{bmatrix} \Phi \tag{10}$$

and a subset of $\mathbb{R}^{m+1}$

$$\mathcal{C} := \{z : z \in \mathbb{R}^{m+1}, z_{m+1} = 1 \text{ and } z_i \geq 0,\ i = 1, 2, \ldots, m\}, \tag{11}$$

respectively. Then a vector x satisfies the two constraints of model (9) if and only if the vector Bx lies in the set $\mathcal{C}$. Hence, model (9) can be rewritten as

$$\min\{\|x\|_0 + \iota_{\mathcal{C}}(Bx) : x \in \mathbb{R}^n\}. \tag{12}$$

Problem (12) is known to be NP-complete due to the non-convexity of the ℓ_0-norm. Thus, there is a need for an algorithm that can pick the sparsest vector x satisfying the relation $Bx \in \mathcal{C}$. To attack this ℓ_0-norm optimization problem, a common approach that appeared in recent literature is to approximate the ℓ_0-norm by its computationally feasible approximations. In the context of compressed sensing, we review several popular choices for defining the ℓ_0-norm as the limit of a sequence. More precisely, for a positive number $\epsilon \in (0,1)$, we consider separable concave functions of the form

$$F_\epsilon(x) := \sum_{i=1}^{n} f_\epsilon(|x_i|), \quad x \in \mathbb{R}^n, \tag{13}$$

where $f_\epsilon : \mathbb{R}_+ \to \mathbb{R}$ is strictly increasing, concave, and twice continuously differentiable such that

$$\lim_{\epsilon \to 0+} F_\epsilon(x) = \|x\|_0, \quad \text{for all} \quad x \in \mathbb{R}^n. \tag{14}$$

Since the function f_ϵ is concave and smooth on $\mathbb{R}_+ := [0, \infty)$, it can be majorized by a simple function formed by its first-order Taylor series expansion at a arbitrary point. Therefore, at any point $v \in \mathbb{R}^n$ the following inequality holds

$$F_\epsilon(x) < F_\epsilon(v) + \langle \nabla F_\epsilon(|v|), |x| - |v| \rangle \tag{15}$$

for all $x \in \mathbb{R}^n$ with $|x| \neq |v|$. Here, for a vector u, we use $|u|$ to denote a vector such that each element of $|u|$ is the absolute value of the corresponding element of u. Clearly, when v is close enough to x, the expression on the right-hand side of (15) provides a reasonable approximation to the one on its left-hand side. Therefore, it is considered as a computationally feasible approximation to the ℓ_0-norm of x. With such an approximation, a simplified problem is solved and its solution is used to formulate a simplified problem which is closer to the ideal problem (12). This process is then repeated until the solutions to the simplified problems become stationary or meet a termination criteria. This procedure is summarized in Algorithm 1.

The constant terms $F_\epsilon(|x^{(k)}|)$ and $\langle \nabla F_\epsilon(|x^{(k)}|), |x^{(k)}| \rangle$ appeared in the optimization problem in Algorithm 1 can be ignored because they are irrelevant to the optimization problem. Hence the expression for $x^{(k+1)}$ in Algorithm 1 can be simplified as

$$x^{(k+1)} \in \operatorname{argmin} \left\{ \langle \nabla F_\epsilon(|x^{(k)}|), |x| \rangle + \iota_{\mathcal{C}}(Bx) : x \in \mathbb{R}^n \right\}. \tag{16}$$

Since f_ϵ is strictly concave and increasing on $\mathbb{R}_+$, f'_ϵ is positive on $\mathbb{R}_+$. Hence, $\langle \nabla F_\epsilon(|x^{(k)}|), |x| \rangle = \sum_{i=1}^{n} f'_\epsilon(|x_i^{(k)}|)|x_i|$ can be viewed as the weighted ℓ_1-norm of x having $f'_\epsilon(|x_i^{(k)}|)$ as its ith weight.

Algorithm 1 (Iterative scheme for model (12))

```
Initialization: choose $\epsilon \in (0,1)$ and let $x^{(0)} \in \mathbb{R}^n$ be an initial point.
repeat($k \geq 0$)
    Step 1: Compute $x^{(k+1)}$:
        $x^{(k+1)} \in \operatorname{argmin} \left\{ F_\epsilon(|x^{(k)}|) + \langle \nabla F_\epsilon(|x^{(k)}|), |x| - |x^{(k)}| \rangle + \iota_{\mathcal{C}}(Bx) : x \in \mathbb{R}^n \right\}.$
until a given stopping criteria is met
```

Thus, the objective function of the above optimization problem is convex. Details for finding a solution to the problem will be presented in the next section.

In the rest of this section, we list several possible choices of the functions in (13) including but not limited to the Mangasarian function in [15], the Geman-McClure function in [9], and the Log-Det function in [8].

The Mangasarian function is given as follows:

$$F_\epsilon(x) = \sum_{i=1}^{n} \left(1 - e^{-|x_i|/\epsilon}\right), \tag{17}$$

where $x \in \mathbb{R}^n$. This function is used to approximate the ℓ_0-norm to obtain minimum-support solutions (that is, solutions with as many components equal to zero as possible). The usefulness of the Mangasarian function was demonstrated in finding sparse solutions of underdetermined linear systems (see [12]).

The Geman-McClure function is given by

$$F_\epsilon(x) = \sum_{i=1}^{n} \frac{|x_i|/\epsilon}{|x_i|/\epsilon + 1}, \tag{18}$$

where $x \in \mathbb{R}^n$. In a regularization framework for recovering an image/signal, its penalty term is often formed by applying this function to a linear transform of the underlying image/signal. This results in strongly homogeneous zones in the solution. The corresponding recovery problem can be recovered from noisy data and preserved intact from small variations of the data [18].

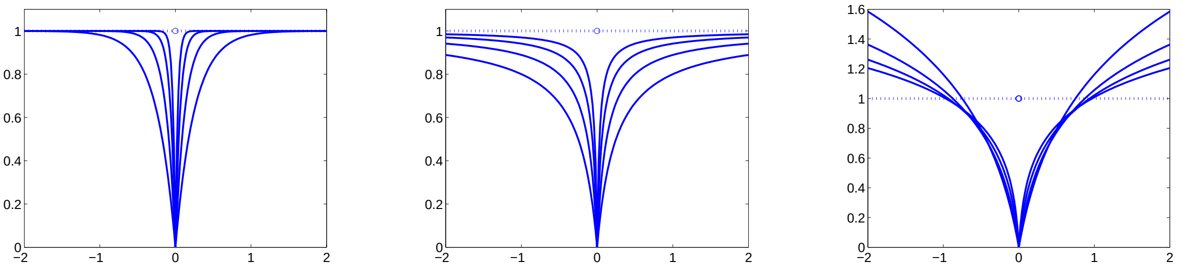
The Log-Det function is defined as

$$F_\epsilon(x) = \sum_{i=1}^{n} \frac{\log(|x_i|/\epsilon + 1)}{\log(1/\epsilon)}, \tag{19}$$

where $x \in \mathbb{R}^n$. Notice that $\|x\|_0$ is equal to the rank of the diagonal matrix $\operatorname{diag}(x)$. The function $F_\epsilon(x)$ is equal to $(\log(1/\epsilon))^{-1} \log(\det(\operatorname{diag}(x) + \epsilon I)) + n$, the logarithm of the determinant of the matrix $\operatorname{diag}(x) + \epsilon I$. Hence, it was named as the Log-Det heuristic and used for minimizing the rank of a positive semidefinite matrix over a convex set in [8]. Constant terms can be ignored since they will not affect the solution of the optimization problem (16). Hence the Log-Det function in in (19) can be replaced by

$$F_\epsilon(x) = \sum_{i=1}^{n} \log(|x_i| + \epsilon). \tag{20}$$

The function F_ϵ for the above three choices are plotted in Figure 3.1 for $n = 1$ and ϵ being $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, and $\frac{1}{32}$. We can see that for a fixed $\epsilon \in (0,1)$ the Mangasarian function is the one which is the most closest to the ℓ_0-norm.



(a) Mangasarian (b) Geman-McClure (c) Log-Det

Figure 3.1: Plots of $F_{\frac{1}{4}}$, $F_{\frac{1}{8}}$, $F_{\frac{1}{16}}$, $F_{\frac{1}{32}}$ with $n = 1$ for (a) the Mangasarian function; (b) the Geman-McClure function; and (c) the Log-Det function.

We point it out that both the Mangasarian function and the Geman-McClure function are bounded by 1, therefore, are non-coercive while the Log-Det function is coercive. This makes difference in convergence analysis of the associated Algorithm 1 that will be presented in the next section. In what follows, the function F_ϵ is the Mangasarian function, the Geman-McClure, or the Log-Det function. We specify it only when it is noted.

4 Convergence Analysis

In this section, we shall give convergence analysis for Algorithm 1. We begin with presenting the following result.

Theorem 4.1 *Given $\epsilon \in (0, 1)$, $x^{(0)} \in \mathbb{R}^n$, and the set $\mathcal{C}$ defined by (11), let the sequence $\{x^{(k)} : k \in \mathbb{N}\}$ be generated by Algorithm 1, where $\mathbb{N}$ is the set of all natural numbers. Then the following three statements hold:*

(i) The sequence $\{F_\epsilon(x^{(k)}) : k \in \mathbb{N}\}$ converges when F_ϵ is corresponding to the Mangasarian function (17)*, the Geman-McClure function* (18) *or the Log-Det function* (20)*;*

(ii) The sequence $\{x^{(k)} : k \in \mathbb{N}\}$ is bounded when F_ϵ is the Log-Det function;

(iii) $\sum_{k=1}^{+\infty} \left\| |x^{(k+1)}| - |x^{(k)}| \right\|_2^2$ is convergent when the sequence $\{x^{(k)} : k \in \mathbb{N}\}$ is bounded.

Proof: We first prove Item (i). The key step for proving it is to show that the sequence $\{F_\epsilon(x^{(k)}) : k \in \mathbb{N}\}$ is decreasing and bounded below. The boundedness of the sequence is due to the fact that $F_\epsilon(0) \leq F_\epsilon(x^{(k)})$. From Step 1 of Algorithm 1 or equation (16), one can immediately have that

$$\iota_{\mathcal{C}}(Bx^{(k+1)}) = 0 \quad \text{and} \quad \langle \nabla F_\epsilon(|x^{(k)}|), |x^{(k+1)}| \rangle \leq \langle \nabla F_\epsilon(|x^{(k)}|), |x^{(k)}| \rangle. \tag{21}$$

By identifying $x^{(k)}$ and $x^{(k+1)}$, respectively, as v and x in (15) and using the inequality in (21), we get $F_\epsilon(x^{(k+1)}) \leq F_\epsilon(x^{(k)})$. Hence, the sequence $\{F_\epsilon(x^{(k)}) : k \in \mathbb{N}\}$ is decreasing and bounded below. Item (i) follows immediately.

When F_ϵ is chosen as the Log-Det function, the coerciveness of F_ϵ together with Item (i) implies that the sequence $\{x^{(k)} : k \in \mathbb{N}\}$ must be bounded, that is, Item (ii) holds.

Finally, we prove Item (iii). From the second-order Taylor expansion of the function F_ϵ at $x^{(k)}$ we have that

$$F_\epsilon(x^{(k+1)}) = F_\epsilon(x^{(k)}) + \langle \nabla F_\epsilon(|x^{(k)}|), |x^{(k+1)}| - |x^{(k)}| \rangle + \frac{1}{2}(|x^{(k+1)}| - |x^{(k)}|)^\top \nabla^2 F_\epsilon(v)(|x^{(k+1)}| - |x^{(k)}|), \tag{22}$$

where v is some point in the line segment linking the points $|x^{(k+1)}|$ and $|x^{(k)}|$ and $\nabla^2 F_\epsilon(v)$ is the Hessian matrix of F_ϵ at the point v.

By (21), the second term on the right-hand of equation (22) is less than 0. By equation (20), $\nabla^2 F_\epsilon(v)$ for v lying in the first octant of $\mathbb{R}^n$ is a diagonal matrix and is equal to

$$-\frac{1}{\epsilon^2}\begin{bmatrix} e^{-\frac{v_1}{\epsilon}} & & \\ & \ddots & \\ & & e^{-\frac{v_n}{\epsilon}} \end{bmatrix}, \quad -2\epsilon \begin{bmatrix} \frac{1}{(v_1+\epsilon)^3} & & \\ & \ddots & \\ & & \frac{1}{(v_n+\epsilon)^3} \end{bmatrix}, \quad \text{or} \quad -\begin{bmatrix} \frac{1}{(v_1+\epsilon)^2} & & \\ & \ddots & \\ & & \frac{1}{(v_n+\epsilon)^2} \end{bmatrix}$$

which corresponds to F_ϵ being the Mangasarian, the Geman-McClure, or the Log-Det function. Hence, the matrix $\nabla^2 F_\epsilon(v)$ is negative definite. Since the sequence $\{x^{(k)} : k \in \mathbb{N}\}$ is bounded, there exists a constant $\rho > 0$ such that

$$\frac{1}{2}(|x^{(k+1)}| - |x^{(k)}|)^\top \nabla^2 F_\epsilon(v)(|x^{(k+1)}| - |x^{(k)}|) \leq -\frac{\rho}{2}\left\| |x^{(k+1)}| - |x^{(k)}| \right\|_2^2.$$

Putting all above results together into (22), we have that

$$F_\epsilon(x^{(k+1)}) \leq F_\epsilon(x^{(k)}) - \frac{\rho}{2}\left\| |x^{(k+1)}| - |x^{(k)}| \right\|_2^2.$$

Summing the above inequality from $k = 1$ to $+\infty$ and using Item (i) we get the proof of Item (iii). □

From Item (iii) of Theorem 4.1, we have $\left\| |x^{(k+1)}| - |x^{(k)}| \right\|_2 \to 0$ as $k \to \infty$.

To further study properties of the sequence $\{x^{(k)} : k \in \mathbb{N}\}$ generated by Algorithm 1, the matrix $B^\top$ is required to have the range space property (RSP) which is originally introduced in [22]. With this property and motivated by the work in [22] we prove that Algorithm 1 can yield a sparse solution for model (12).

Prior to presenting the definition of the RSP, we introduce the notation to be used throughout the rest of this paper. Given a set $S \subset \{1, 2, \ldots, n\}$, the symbol $|S|$ denotes the cardinality of S, and $S^c := \{1, 2, \ldots, n\} \setminus S$ is the complement of S. Recall that for a vector u, by abuse of notation, we also use $|u|$ to denote the vector whose elements are the absolute values of the corresponding elements of u. For a given matrix A having n columns, a vector u in $\mathbb{R}^n$, and a set $S \subset \{1, 2, \ldots, n\}$, we use the notation A_S to denote the submatrix extracted from A with column indices in S, and u_S the subvector extracted from u with component indices in S.

Definition 4.2 (Range Space Property (RSP)) *Let A be an $m \times n$ matrix. Its transpose $A^\top$ is said to satisfy the* range space property (RSP) *of order K with a constant $\rho > 0$ if for all sets $S \subseteq \{1, \ldots, n\}$ with $|S| \geq K$ and for all ξ in the range space of $A^\top$ the following inequality holds*

$$\|\xi_{S^c}\|_1 \leq \rho \|\xi_S\|_1.$$

We remark that if the transpose of an $m \times n$ matrix B has the RSP of order K with a constant $\rho > 0$, then for every non-empty set $S \subseteq \{1, \ldots, n\}$, the transpose of the matrix B_S, denoted by $B_S^\top$, has the RSP of order K with constant ρ as well.

The next result shows that if the transpose of the matrix B in Algorithm 1 possesses the RSP, then Algorithm 1 can lead to a sparse solution for model (12). To this end, we define a mapping $\sigma : \mathbb{R}^d \to \mathbb{R}^d$ such that the ith component of the vector $\sigma(u)$ is the ith largest component of $|u|$.

Proposition 4.3 *Let B be the $(m+1) \times n$ matrix be defined by* (10) *and let $\{x^{(k)} : k \in \mathbb{N}\}$ be the sequence generated by Algorithm 1. Assume that the matrix $B^\top$ has the RSP of order K with $\rho > 0$ satisfying $(1+\rho)K < n$. Suppose that the sequence $\{x^{(k)} : k \in \mathbb{N}\}$ is bounded. Then $(\sigma(x^{(k)}))_n$ the nth largest component of $x^{(k)}$ converges to* 0.

Proof: Suppose this proposition is *false.* Then there exist a constant $\gamma > 0$ and a subsequence $\{x^{(k_j)} : j \in \mathbb{N}\}$ such that $(\sigma(x^{(k_j)}))_n \geq 2\gamma > 0$ for all $j \in \mathbb{N}$. From Item (iii) of Theorem 4.1 we have that

$$(\sigma(x^{(k_j+1)}))_n \geq \gamma \tag{23}$$

for all sufficient large j. For simplicity, we set $y^{(k_j)} := \nabla F_\epsilon(|x^{(k_j)}|)$. Hence, by inequality (23) and F_ϵ, we know that

$$|x^{(k_j)}| > 0 \quad |x^{(k_j+1)}| > 0, \quad \text{and} \quad y^{(k_j)} > 0 \tag{24}$$

for all sufficient large j. In what follows, we assume that the integer j is large enough such that the above inequalities in (24) hold.

Since the vector $x^{(k_j+1)}$ is obtained through Step 1 of Algorithm 1, i.e., equation (16), then by Fermat's rule and the chain rule of subdifferential we have that

$$0 = \mathrm{diag}(y^{(k_j)})\partial\|\cdot\|_1(\mathrm{diag}(y^{(k_j)})x^{(k_j+1)}) + B^\top b^{(k_j+1)},$$

where $b^{(k_j+1)} \in \partial\iota_C(Bx^{(k_j+1)})$. By (24), we get

$$\partial\|\cdot\|_1(\mathrm{diag}(y^{(k_j)})x^{(k_j+1)}) = \{(\mathrm{sign}(x_1^{(k_j+1)}), \mathrm{sign}(x_2^{(k_j+1)}), \ldots, \mathrm{sign}(x_n^{(k_j+1)}))^\top\}.$$

Thus

$$y^{(k_j)} = |\xi^{(k_j+1)}|,$$

where $\xi^{(k_j+1)} = B^\top b^{(k_j+1)}$ is in the range of $B^\top$.

Let S be the set of indices corresponding to the K smallest components of $|\xi^{(k_j+1)}|$. Hence,

$$\sum_{i=1}^{n-K}(\sigma(y^{(k_j)}))_i = \|\xi_{S^c}^{(k_j+1)}\|_1 \quad \text{and} \quad \sum_{i=n-K+1}^{n}(\sigma(y^{(k_j)}))_i = \|\xi_S^{(k_j+1)}\|_1$$

Since $B^\top$ has the RSP of order K with the constant ρ, we have that $\|\xi_{S^c}^{(k_j+1)}\|_1 \leq \rho\|\xi_S^{(k_j+1)}\|_1$. Therefore,

$$\sum_{i=1}^{n-K}(\sigma(y^{(k_j)}))_i \leq \rho \sum_{i=n-K+1}^{n}(\sigma(y^{(k_j)}))_i. \tag{25}$$

However, by the definition of σ, we have that

$$\sum_{i=1}^{n-K}(\sigma(y^{(k_j)}))_i \geq (n-K)(\sigma(y^{(k_j)}))_{n-K+1} \quad \text{and} \quad \sum_{i=n-K+1}^{n}(\sigma(y^{(k_j)}))_i \leq K(\sigma(y^{(k_j)}))_{n-K+1}.$$

These inequalities together with the condition $(1+\rho)K < n$ lead to

$$\sum_{i=1}^{n-K}(\sigma(y^{(k_j)}))_i > \rho \sum_{i=n-K+1}^{n}(\sigma(y^{(k_j)}))_i,$$

which contradicts to (25). This completes the proof of the proposition. □

From Proposition 4.3, we conclude that a sparse solution is guaranteed via Algorithm 1 if the transpose of B satisfies the RSP. Next, we answer how sparse this solution will be. To this end, we introduce some notation and develop a technical lemma. For a vector $x \in \mathbb{R}^d$, we denote by $\tau(x)$ the set of the indices of non-zero elements of x, i.e., $\tau(x) := \{i : x_i \neq 0\}$. For a sequence $\{x^{(k)} : k \in \mathbb{N}\}$, a positive number μ, and an integer k, we define $I_\mu(x^{(k)}) := \{i : |x_i^{(k)}| \geq \mu\}$.

Lemma 4.4 *Let B be the $(m+1)\times n$ matrix defined by* (10)*, let F_ϵ be the Log-Det function defined by* (20)*, and let $\{x^{(k)} : k \in \mathbb{N}\}$ be the sequence generated by Algorithm 1. Assume that the matrix $B^\top$ has the RSP of order K with $\rho > 0$ satisfying $(1+\rho)K < n$. If there exist $\mu > \rho\epsilon n$ such that $|I_\mu(x^{(k)})| \geq K$ for all sufficient large k, then there exists a $k'' \in \mathbb{N}$ such that $\|x^{(k)}\|_0 < n$ and $\tau(x^{(k+1)}) \subseteq \tau(x^{(k'')})$ for all $k > k''$.*

Proof: Set $y^{(k)} := \nabla F_\epsilon(|x^{(k)}|)$. Since $x^{(k+1)}$ is a solution to the optimization problem (16), then by Fermat's rule and the chain rule of subdifferential we have that

$$0 \in \operatorname{diag}(y^{(k)})\partial\|\cdot\|_1(\operatorname{diag}(y^{(k)})x^{(k+1)}) + B^\top b^{(k+1)},$$

where $b^{(k+1)} \in \partial\iota_C(Bx^{(k+1)})$. Hence, if $x_i^{(k+1)} \neq 0$, we have that $y_i^{(k)} = |(B^\top b^{(k+1)})_i|$.

For $i \in I_\mu(x^{(k)})$, we have that $|x_i^{(k)}| \geq \mu$ and $y_i^{(k)} = f'_\epsilon(|x_i^{(k)}|) \leq f'_\epsilon(\mu)$ for all $k \in \mathbb{N}$, where $f_\epsilon = \log(\cdot + \epsilon)$. Furthermore, there exist a k' such that $|x_i^{k+1}| > 0$ for $i \in I_\mu(x^{(k)})$ and $k \geq k'$ due to Item (iii) in Theorem 4.1. Thus, we have for all $k \geq k'$

$$\sum_{i\in I_\mu(x^{(k)})} |(B^\top b^{(k+1)})_i| = \sum_{i\in I_\mu(x^{(k)})} y_i^{(k)} \leq \sum_{i\in I_\mu(x^{(k)})} f'_\epsilon(\mu) \leq W^*,$$

where $W^* = n\lim_{\epsilon\to 0+} f'_\epsilon(\mu) = \frac{n}{\mu}$ is a positive number dependent on μ.

Now, we are ready to prove $\|x^{(k)}\|_0 < n$ for all $k > k''$. By Proposition 4.3, we have that $(\sigma(x^{(k)}))_n \to 0$ when $k \to +\infty$. Therefore, there exists an integer $k'' > k'$ such that $|I_\mu(x^{(k)})| \geq K$ and $0 \leq \sigma(x^{(k)}))_n < \min\{\frac{\mu}{\rho n} - \epsilon, \mu\}$ for all $k \geq k''$. Let i_0 be the index such that $|x_{i_0}^{(k'')}| = (\sigma(x^{(k'')}))_n$. We will show that $x_{i_0}^{(k''+1)} = 0$. If this statement is not true, that is, $x_{i_0}^{(k''+1)}$ is not zero, then

$$|(B^\top b^{(k''+1)})_{i_0}| = f'_\epsilon(|x_{i_0}^{(k'')}|) = \frac{1}{|x_{i_0}^{(k'')}| + \epsilon} > \rho W^*. \tag{26}$$

However, since i_0 is not in the set $I_\mu(x^{(k'')})$ and $B^\top$ satisfies the RSP, we have that

$$|(B^\top b^{(k''+1)})_{i_0}| \leq \sum_{i\notin I_\mu(x^{(k'')})} |(B^\top b^{(k''+1)})_i| \leq \rho \sum_{i\in I_\mu(x^{(k'')})} |(B^\top b^{(k''+1)})_i| \leq \rho W^*,$$

which contradicts to (26). Hence, we have that $x_{i_0}^{(k''+1)} = 0$ and $|\tau(x^{(k''+1)})| < n$. By replacing k'' by $k''+1$ and repeating this process we can obtain $x_{i_0}^{(k''+\ell)} = 0$ for all $\ell \in \mathbb{N}$. Therefore, $\|x\|_0 < n$ for all $k > k''$. This process can be also applied to other components satisfying $x_i^{(k''+1)} = 0$. Thus there exists a $k'' \in \mathbb{N}$ such that $\tau(x^{(k)}) \subseteq \tau(x^{(k'')})$ for all $k \geq k''$. □

With Lemma 4.4, the next result shows that when the transpose of B satisfies the RSP there exists a cluster point of the sequence generated by Algorithm 1 that is sparse and satisfies the consistency condition.

Theorem 4.5 *Let B be the $(m+1)\times n$ matrix defined by (10), let F_ϵ be the Log-Det function defined by (20), and let $\{x^{(k)}: k\in\mathbb{N}\}$ be the sequence generated by Algorithm 1. Assume that the matrix $B^\top$ has the RSP of order K with $\rho>0$ satisfying $(1+\rho)K<n$. Then there is a subsequence $\{x^{(k_j)}: j\in\mathbb{N}\}$ that converges to a $\lfloor(1+\rho)K\rfloor$-sparse solution, that is $(\sigma(x^{(k_j)}))_{\lfloor(1+\rho)K+1\rfloor}\to 0$ as $j\to+\infty$ and $\epsilon\to 0$.*

Proof: Suppose the theorem is *false*. Then there exist μ^*, for any $0<\epsilon^*<\frac{\mu^*}{\rho n}$, there exist a $\epsilon\in(0,\epsilon^*)$ and k' such that $(\sigma(x^{(k)}))_{\lfloor(1+\rho)K+1\rfloor}\geq\mu^*$ for all $k\geq k'$. It implies that for all $k\geq k'$

$$|I_{\mu^*}(x^{(k)})|\geq\lfloor(1+\rho)K+1\rfloor>(1+\rho)K>K. \tag{27}$$

By Lemma 4.4, there exist a $k''\geq k'$ such that $\|x^{(k)}\|_0<n$ and $\tau(x^{(k+1)})\subseteq\tau(x^{(k'')})$ for all $k\geq k''$. Let $S=\tau(x^{(k'')})$. Thus $x^{(k)}_{S^c}=0$ for all $k\geq k''$. Therefore, the optimization problem (16) for updating $x^{(k+1)}$ can be reduced to the following one

$$x^{k+1}_S\in\arg\min\{\langle(\nabla F_\epsilon(|x^{(k)}|))_S,u\rangle+\iota((B_S)u): u\in\mathbb{R}^{|S|}\}. \tag{28}$$

If $|\tau(x^{(k'')})|>|I_{\mu^*}(x^{(k'')})|$, from (27) we have $(1+\rho)K<|S|$. Thus from Lemma 4.4 and $B_S^\top$ having RSP with the same parameters, there exist a $k'''>k''$ such that $\tau(x^{(k)})<\tau(x^{(k'')})$ for all $k\geq k'''$. Therefore, by induction, there must exist a $\tilde{k}$ such that for all $k\geq\tilde{k}$

$$\tau(x^{(k)})=I_{\mu^*}(x^{(k)}),\ \tau(x^k)\subseteq\tau(x^{(\tilde{k})}).$$

It means that for all $k\geq\tilde{k}$ all the nonzero components of $x^{(k)}$ are bounded below by μ^*. Therefore, for any $k\geq\tilde{k}$, the updating equation (16) is reduced by (28) with $S=I_{\mu^*}(x^{(k)})$. From Lemma 4.3 we get $[\sigma(x^{(k)})]_{|S|}\to 0$ which contradicts with $|x^k_{|S|}|\geq\mu^*$. Therefore, we get this theorem. □

5 An Implementation of Algorithm 1

In this section, we describe in detail an implementation of Algorithm 1 and show how to select the parameters of the associated algorithm.

Solving problem (16) is the main issue for Algorithm 1. A general model related to (16) is

$$\min\{\|\Gamma x\|_1+\varphi(Bx): x\in\mathbb{R}^n\}, \tag{29}$$

where Γ is a diagonal matrix with positive diagonal elements and φ is in $\Gamma_0(\mathbb{R}^{m+1})$. In particular, if we choose $\Gamma=\nabla F_\epsilon(|x^{(k)}|)$ and $\varphi=\iota_{\mathcal{C}}$, where $x^{(k)}$ is a vector in $\mathbb{R}^n$, ϵ is a positive number, $\mathcal{C}$ is given by (11), and F_ϵ is a function given by (13), then model (29) reduces to the optimization problem in Algorithm 1.

We solve model (29) by using recently developed first-order primal-dual algorithm (see, e.g., [6, 14]). To present this algorithm, we need two concepts in convex analysis, namely, the proximity operator and conjugate function. The proximity operator was introduced in [16, 17]. For a function $f\in\Gamma_0(\mathbb{R}^d)$, the proximity operator of f with parameter λ, denoted by $\mathrm{prox}_{\lambda f}$, is a mapping from $\mathbb{R}^d$ to itself, defined for a given point $x\in\mathbb{R}^d$ by

$$\mathrm{prox}_{\lambda f}(x):=\mathrm{argmin}\left\{\frac{1}{2\lambda}\|u-x\|_2^2+f(u): u\in\mathbb{R}^d\right\}.$$

Algorithm 2 PD-subroutine (The first-order primal-dual algorithm for solving (29))

```
**Input**: the $(m+1)\times n$ matrix $B$ defined by (10); two positive numbers $\alpha$ and $\beta$ satisfying the relation $\alpha\beta < \frac{1}{\|B\|^2}$; the $n \times n$ diagonal matrix $\Gamma$ with all diagonal elements positive; and the function $\varphi \in \Gamma_0(\mathbb{R}^n)$.
**Initialization**: $i = 0$ and an initial guess $(u^{-1}, u^0, x^0) \in \mathbb{R}^{m+1} \times \mathbb{R}^{m+1} \times \mathbb{R}^n$
**repeat**$(i \geq 0)$
    Step 1: Compute $x^{i+1}$:
        $$x^{i+1} = \mathrm{prox}_{\alpha\|\cdot\|_1\circ\Gamma}\left(x^i - \alpha B^\top(2u^i - u^{i-1})\right)$$
    Step 2: Compute $u^{i+1}$:
        $$u^{i+1} = \mathrm{prox}_{\beta\varphi^*}(u^i + \beta Bx^{i+1})$$
    Step 3: Set $i := i+1$.
**until** a given stopping criteria is met and the corresponding vectors $u^i$, $u^{i+1}$, and $x^{i+1}$ are denoted by $u^{cur}$, $u^{new}$, and $x^{new}$, respectively.
**Output**: $(u^{cur}, u^{new}, x^{new}) = \mathrm{PD}(\alpha, \beta, B, \Gamma, \varphi, u^{-1}, u^0, x^0)$
```

The conjugate of $f \in \Gamma_0(\mathbb{R}^d)$ is the function $f^* \in \Gamma_0(\mathbb{R}^d)$ defined at $z \in \mathbb{R}^d$ by

$$f^*(z) := \sup\{\langle x, z\rangle - f(x) : x \in \mathbb{R}^d\}.$$

With these notation, the first-order primal-dual (PD) method for solving (29) is summarized in Algorithm 2 (referred to as PD-subroutine).

Theorem 5.1 *Let B be an $(m+1)\times n$ matrix defined by* (10)*, let $\mathcal{C}$ be the set given by* (11)*, let α and β be two positive numbers, and let L be a positive such that $L \geq \|B\|^2$, where $\|B\|$ is the largest singular value of B. If*

$$\alpha\beta L < 1,$$

then for any arbitrary initial vector $(x^{-1}, x^0, u^0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{m+1}$, the sequence $\{x^k : k \in \mathbb{N}\}$ generated by Algorithm 2 converges to a solution of model (29)*.*

The proof of Theorem 5.1 follows immediately from Theorem 1 in [6] or Theorem 3.5 in [14]. We skip its proof here.

Both proximity operators $\mathrm{prox}_{\alpha\|\cdot\|_1\circ\Gamma}$ and $\mathrm{prox}_{\beta\varphi^*}$ should be computed easily and efficiently in order to make the iterative scheme in Algorithm 2 numerically efficient. Indeed, the proximity operator $\mathrm{prox}_{\alpha\|\cdot\|_1\circ\Gamma}$ is given at $z \in \mathbb{R}^n$ as follows: for $j = 1, 2, \ldots, n$

$$\left(\mathrm{prox}_{\alpha\|\cdot\|_1\circ\Gamma}(z)\right)_j = \max\{|z_j| - \alpha\gamma_j, 0\}\cdot \mathrm{sign}(z_j), \tag{30}$$

where γ_j is the jth diagonal element of Γ. Using the well-known Moreau decomposition (see, e.g. [1, 16])

$$\mathrm{prox}_{\beta\varphi^*} = I - \beta\ \mathrm{prox}_{\frac{1}{\beta}\varphi} \circ \left(\frac{1}{\beta}I\right), \tag{31}$$

we can compute the proximity operator $\mathrm{prox}_{\beta\varphi^*}$ via $\mathrm{prox}_{\frac{1}{\beta}\varphi}$ which depends on a particular form of the function φ. As our purpose is to develop algorithms for the optimization problem in Algorithm 1, we need to compute the proximity operator of $\iota_{\mathcal{C}}^*$ which is given in the following.

Lemma 5.2 *If $\mathcal{C}$ is the set given by* (11) *and β is a positive number, then for $z \in \mathbb{R}^{m+1}$ we have that*

$$\mathrm{prox}_{\beta\iota_{\mathcal{C}}^*}(z) = (z_1 - (z_1)_+, z_2 - (z_2)_+, \dots, z_m - (z_m)_+, z_{m+1} - \beta), \tag{32}$$

where $(s)_+$ is s if $s \geq 0$ and 0 otherwise.

Proof: We first give an explicit form for the proximity operator $\mathrm{prox}_{\frac{1}{\beta}\iota_{\mathcal{C}}}$. Note that $\iota_{\mathcal{C}} = \frac{1}{\beta}\iota_{\mathcal{C}}$ for $\beta > 0$ and $\iota_{\mathcal{C}}(z) = \iota_{\{1\}}(z_{m+1}) + \sum_{i=1}^m \iota_{[0,\infty)}(z_i)$, for $z \in \mathbb{R}^{m+1}$. Hence, we have that

$$\mathrm{prox}_{\frac{1}{\beta}\iota_{\mathcal{C}}}(z) = ((z_1)_+, (z_2)_+, \dots, (z_m)_+, 1), \tag{33}$$

where $(s)_+$ is s if $s \geq 0$ and 0 otherwise. Here we use the facts that $\mathrm{prox}_{\iota_{[0,+\infty)}}(s) = (s)_+$ and $\mathrm{prox}_{\iota_{\{1\}}}(s) = 1$ for any $s \in \mathbb{R}$.

By the Moreau decomposition (31), we have that $\mathrm{prox}_{\beta\iota_{\mathcal{C}}^*}(z) = z - \beta\mathrm{prox}_{\frac{1}{\beta}\iota_{\mathcal{C}}}(\frac{1}{\beta}z)$. This together with equation (33) yields (32). □

Next, we comment on the diagonal matrix Γ in model (29). When the function φ in model (29) is chosen to be ι_C, then the relation $a\varphi = \varphi$ holds for any positive number a. Hence, by rescaling the diagonal matrix Γ in model (29) with any positive number, that does not alter the solutions of model (29). Therefore, we can assume that the largest diagonal entry of Γ is always equal to one.

In applications of Theorem 5.1 as in Algorithm 2, we should make the product of α and β as close to $1/\|B\|^2$ as possible. In our numerical simulations, we always set

$$\alpha = \frac{0.999}{\beta\|B\|^2}. \tag{34}$$

In such the way, β is essentially the only parameter that needs to be determined.

Prior to computing α for a given β by equation (34), we need to know the norm of the matrix B. When $\min\{m, n\}$ is small, the norm of the matrix B can be computed directly. When $\min\{m, n\}$ is large, an upper bound of the norm of the matrix B is estimated in terms of the size of B as follows.

Proposition 5.3 *Let Φ be an $m \times n$ matrix with i.i.d. standard Gaussian entries and y be an m-dimensional vector with its component being $+1$ or -1. We define an $(m+1) \times n$ matrix B from Φ and y via equation* (10). *Then*

$$\mathbb{E}\{\|B\|\} \leq \sqrt{m+1}(\sqrt{n} + \sqrt{m}).$$

Moreover,

$$\|B\| \leq \sqrt{m+1}(\sqrt{n} + \sqrt{m} + t)$$

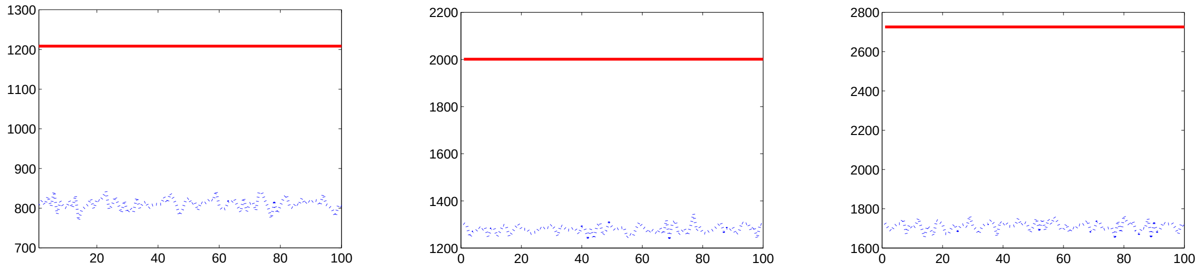
holds with probability at least $1 - 2e^{-t^2/2}$ for all $t \geq 0$.

Proof: By the structure of the matrix B in (10), we know that

$$\|B\| \leq \left\| \begin{bmatrix} \mathrm{diag}(y) \\ y^\top \end{bmatrix} \right\| \cdot \|\Phi\|.$$

Therefore, we just need to compute the norms on the right-hand side of the above inequality. Denote by I_m the $m \times m$ identity matrix and 1_m the vector with all its components being 1. Then

$$\begin{bmatrix} \mathrm{diag}(y) \\ y^\top \end{bmatrix} \begin{bmatrix} \mathrm{diag}(y) & y \end{bmatrix} = \begin{bmatrix} I_m & 1_m \\ 1_m^\top & m \end{bmatrix},$$



(a) $(m, n) = (500, 1000)$ (b) $(m, n) = (1000, 1000)$ (c) $(m, n) = (1500, 1000)$

Figure 5.2: Plots (the dotted blue lines) of the norms of the $(m+1) \times n$ matrices B . The heights of the red lines are computed by the formula $\sqrt{m+1}(\sqrt{n}+\sqrt{m})$.

which is a special arrow-head matrix and has $m + 1$ as its largest eigenvalue (see [20]). Hence,

$$\left\| \begin{bmatrix} \mathrm{diag}(y) \\ y^\top \end{bmatrix} \right\| = \sqrt{m+1}.$$

Furthermore, by using random matrix theory for the matrix Φ, we know that $\mathbb{E}\{\|\Phi\|\} \leq \sqrt{n} + \sqrt{m}$ and $\|\Phi\| \leq \sqrt{n} + \sqrt{m} + t$ with probability at least $1 - 2e^{-t^2/2}$ for all $t \geq 0$ (see, e.g., [7]). This completes the proof of this proposition. □

The dotted blue lines in Figure 5.2(a), (b), and (c) depict the norm of B for 100 randomly generated matrices Φ and vectors y for the pair (m, n) with three different choices $(500, 1000)$, $(1000, 1000)$, and $(1500, 1000)$, respectively. Corresponding to these choices, the mean values of $\|B\|$ are about 815, 1276, and 1711 while the upper bounds of the expected values of $\|B\|$ by Proposition 5.3 are about 1208, 2001, and 2726, respectively. We can see that the norm of B varies with its size and turns to be a big number when the value of $\min\{m, n\}$ is relatively large. As a consequence, the parameter α or β must be very small relative to the other by equation (34). Therefore, in what follows, the used matrix B in model (29) is considered to have been rescaled in the following way:

$$\frac{B}{\|B\|} \quad \text{or} \quad \frac{B}{\sqrt{m+1}(\sqrt{n}+\sqrt{m})} \tag{35}$$

when the norm of B can be computed easily or not.

The complete procedure for model (12) and how the PD-subroutine is employed are summarized in Algorithm 3.

6 Numerical Simulations

In this section, we demonstrate the performance of Algorithm 3 for 1-bit compressive sampling reconstruction in terms of accuracy and consistency and compare it with the BIHT algorithm.

Through this section, all random $m \times n$ matrices Φ and length-n, s-sparse vectors x are generated based on the following assumption: entries of Φ and x on its support are i.i.d. Gaussian random variables with zero mean and unit variances. The locations of the nonzero entries (i.e., the support) of x are randomly permuted. We then generate the 1-bit observation vector y by equation (2). We obtain reconstruction of $x^\star$ from y by using the BIHT and Algorithm 3. The quality of the

```
Algorithm 3 (Iterative scheme for model (12))
    Input: the $(m+1)\times n$ matrix $B$ formed by an $m\times n$ matrix $\Phi$ and an $m$-dimensional vector $y$ via (10); the set $\mathcal{C}$ given by (11); $\epsilon \in (0,1)$, and $\tau > 0$; $\alpha_{\max}$ and $\epsilon_{\min}$ be two real numbers; the maximum iteration number $k_{\max}$.
    Initialization: normalizing $B$ according to (35); $\Gamma$ being the $n\times n$ identity matrix; an initial guess $(u^{old_0}, u^{cur_0}, x^{(0)}) \in \mathbb{R}^{m+1}\times\mathbb{R}^{m+1}\times\mathbb{R}^n$; and initial parameters $\beta$ and $\alpha = 0.999/\beta$.
    while $k < k_{\max}$ do
        Step 1: Compute
            $(u^{old_{k+1}}, u^{cur_{k+1}}, x^{(k+1)}) = \mathrm{PD}(\alpha, \beta, B, \Gamma, \iota_{\mathcal{C}}, u^{old_k}, u^{cur_k}, x^{(k)})$
        Step 2: Update $\Gamma$ as the scaled matrix $\mathrm{diag}(\nabla F_\epsilon(x^{(k+1)}))$ such that the largest diagonal element of $\Gamma$ is one.
        Step 3: If $\alpha < \alpha_{\max}$, update $\alpha \leftarrow 2\alpha$,   $\beta \leftarrow \beta/2$; if $\epsilon > \epsilon_{\min}$, update $\epsilon \leftarrow \tau\epsilon$;
        Step 4: Update $k \leftarrow k+1$.
    end while
    Output: $x^{(k_{\max})}$
```

reconstructed $x^\star$ is measured in terms of the signal-to-noise ratio (SNR) in dB

$$\mathrm{SNR}(x, x^\star) = 20\log_{10}\left(\left\|\frac{x}{\|x\|}\right\|_2 \Big/ \left\|\frac{x}{\|x\|} - \frac{x^\star}{\|x^\star\|}\right\|_2\right).$$

The accuracy of the BIHT and Algorithm 3 is measured by the average of SNR values over 100 trials unless otherwise noted. For all figures in this section, results by the BIHT and Algorithm 3 with the Mangasarian function (17), the Geman-McClure function (18) and the Log-Det function (20) are marked by the symbols "▽", "○", "+", and "⋆", respectively.

6.1 Effects of using inaccurate sparsity on the BIHT

The BIHT requires the availability of the sparsity of the underlying signals. This requirement is, however, not known in practical applications. In this subsection, we demonstrate through numerical experiments that the mismatched sparsity for a signal will degenerate the performance of the BIHT.

To this end, we fix $n = 1000$ and $s = 10$ and consider two cases of m being 500 and 1000. For each case, we vary the sparsity input for the BIHT from 8 to 12 in which 10 is the only right choice. Therefore, there are total ten configurations. For each configuration, we record the SNR values and the numbers of sign constraints not being satisfied of the reconstructed signals by the BIHT.

Figure 6.3 depicts the SNR values of the experiments. The plots in the top row of Figure 6.3 are for the case $m = 500$ while the plots in the bottom row are for the case $m = 1000$. The marks in each plot represent the pairs of the SNR values with a mismatched sparsity input (i.e., $s = 8$, $s = 9$, $s = 11$, or $s = 12$ corresponding to the column 1, 2, 3, or 4) and with the correct sparsity input (i.e., $s = 10$). A mark below the red line indicates that the BIHT with the correct sparsity input works better than the one with an incorrect sparsity input. A mark that is far away from the red line indicates the BIHT with the correct sparsity input works much better than the one with an incorrect sparsity input or vice versa. Except the second plot in the top, we can see that the BIHT with the correct sparsity input performs better than the one with an inaccurate sparsity input. In particular, when an underestimated sparsity input to the BIHT is used, the performance of the BIHT will be significantly reduced (see the plots in the first two columns of Figure 6.3).

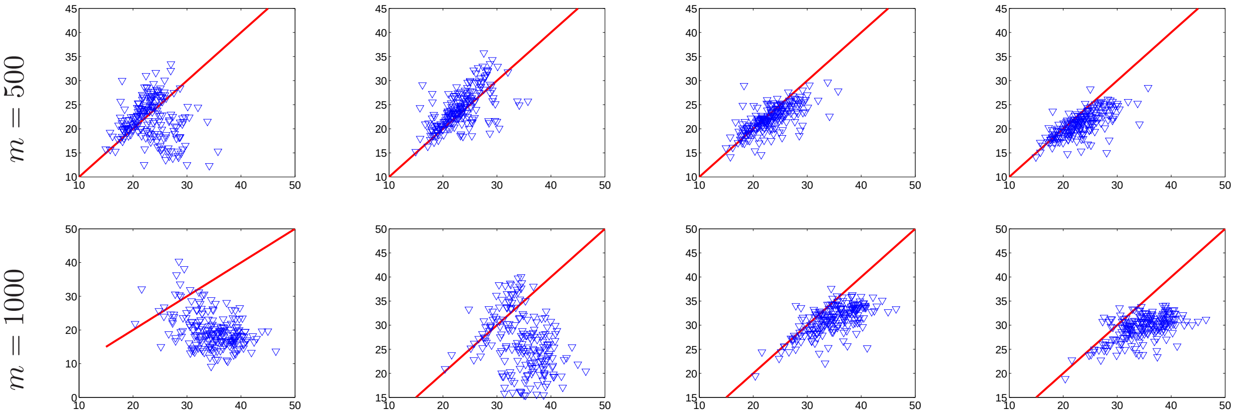
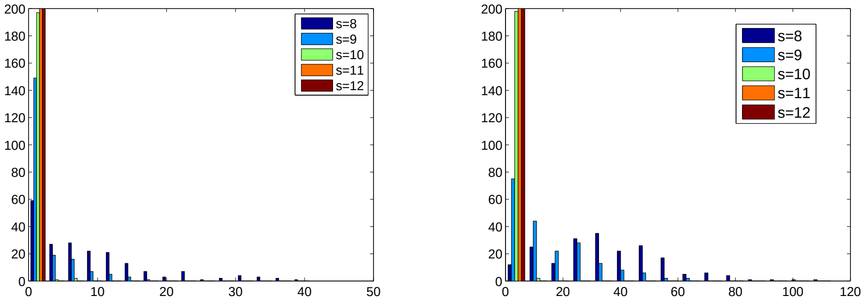


Figure 6.3: The marks in the plots (from the left column to the right column) represent the pairs of the SNR values from the BIHT with the correct sparsity input (i.e., $s = 10$) and incorrect sparsity inputs 8, 9, 11, and 12, respectively. We fix $n = 1000$.



(a) $(m, n) = (500, 1000)$ (b) $(m, n) = (1000, 1000)$

Figure 6.4: The histograms of the numbers of unsatisfied consistency conditions over 200 trials.

When an overestimated sparsity input to the BIHT is used, majority marks are under the red lines and are relatively closer to the red lines than those from the BIHT with underestimated sparsity input. We further report that the average SNR values for the sparsity input $s = 8$, 9, 10, 11, and 12 for $m = 500$ are 21.89dB, 24.18dB, 23.25dB, 22.10dB, and 21.00dB, respectively. Similarly, for $m = 1000$, the average SNR values for the sparsity input $s = 8$, 9, 10, 11, and 12 are 19.77dB, 26.37dB, 34.74dB, 31.12dB, and 29.46dB, respectively.

Figure 6.4 (a) and (b) illustrate the histograms of the numbers of unsatisfied consistency conditions over 200 trials for $m = 500$ and 1000, respectively. We can see from Figure 6.4 (a) that the use of an underestimated sparsity constraint ($s = 8$ or 9) will tend to yield, on average, a solution with a large amount of sign constraints unsatisfied, in other words, under the current setting the solution to model (7) via the BIHT does not satisfy equation (2). As expected, when an overestimated sparsity constraint ($s = 11$ or 12) is used, the sign constraints are usually satisfied.

In summary, we conclude that a proper chosen sparsity constraint is critical for the success of the BIHT.

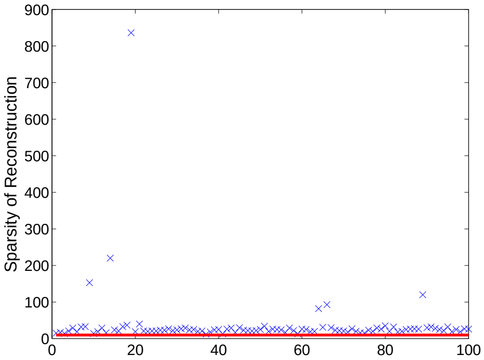


Figure 6.5: Results for Plan-Vershynin's model using Linear programming over 100 trials.

6.2 Plan-Vershynin's model for 1-bit reconstruction

Both our model (9) and Plan-Vershynin's model (8) use the same constraint conditions. Their objective functions are different. Our model uses ℓ_0-norm while Plan-Vershynin's model uses ℓ_1-norm. As suggested in [19], linear programming can be applied for the Plan-Vershynin model. We report here some numerical results for this model.

In our simulations, we fix $n = 1000$, $m = 1000$, and $s = 10$. All simulations were performed 100 trials. Figure 6.5 illustrates the sparsity of the reconstructions of all trials which are clearly greater than 10 (indicated by the solid red line in the figure). The average sparsity of the reconstructions over 100 trials is 23.42. Recall that the average SNR values of all reconstructions by the BIHT is 34.74dB.

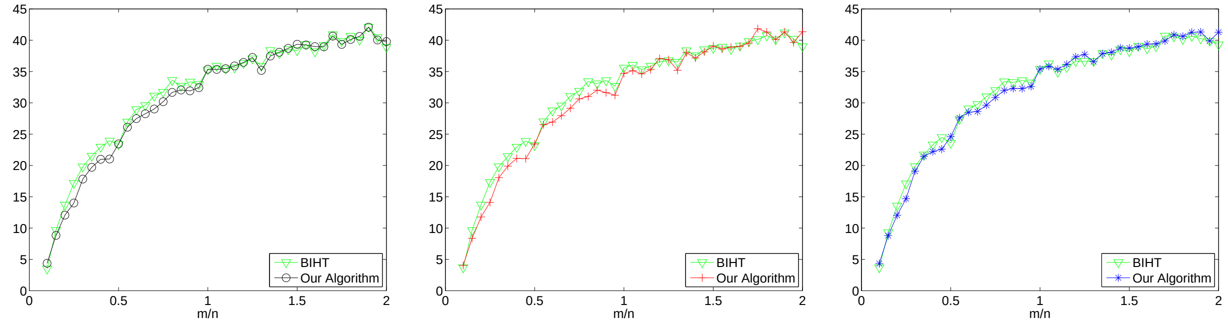
6.3 Performance of Algorithm 3

Prior to applying Algorithm 3 for 1-bit compressive sampling problem, parameters $k_{\max}$, τ, $\alpha_{\max}$, $\epsilon_{\min}$, α, and ϵ in Algorithm 3 need to be determined. Under the aforementioned setting for the random matrix Φ and sparse signal x, we fix $k_{\max} = 17$, $\tau = \frac{1}{2}$, $\alpha_{\max} = 8000$, $\epsilon_{\min} = 10^{-5}$. For the functions F_ϵ defined by (17), (18), and (20), we set the pair of initial parameters (α, ϵ) as $(500, 0.25)$, $(500, 0.125)$, and $(250, 0.125)$, respectively. The iterative process in the PD-subroutine is forced to stop if the corresponding number of iteration exceeds 300. These parameters are used in all simulations performed by Algorithm 3 in the rest of this section.

To evaluate the performance of Algorithm 3 in terms of SNR values at various scenarios, we consider three configurations for the size of the random matrix Φ and the sparsity of the vector x. In the first configuration, we fix $n = 1000$ and $s = 10$ and vary m such that the ratio m/n is between 0.1 and 2. In the second configuration, we fix $m = 1000$ and $n = 1000$ and vary the sparsity of x from 1 to 20. In the third configuration, we fix $m = 1000$ and $s = 10$ and vary n from 500 to 1400.

For every cases in each configuration, we compare the accuracy of Algorithm 3 with the BIHT by computing the average of SNR values over 100 trials. For the given parameters and stopping criteria adopted by Algorithm 3, the estimate $x^{(k_{\max})}$ may not satisfy the consistency condition (3), that is the signs of measurements of the estimate $x^{(k_{\max})}$ are not completely consistent with that of the original measurements. Thus, for a fair comparison, we only compute the average of SNR values for those trials that both reconstructions from the BIHT and Algorithm 3 satisfy the consistency condition (3) and we say the corresponding trials are valid.

For the first configuration, the SNR values in decibels of the average reconstruction errors by both the BIHT and Algorithm 3 are depicted in Figure 6.6. The plots demonstrate that our proposed algorithm performs as equally good as the BIHT, in particular, when m/n is greater than 1, even thought our algorithm does not require to know the exact sparsity of the original signal. We can see that Algorithm 3 with the Log-Det function (20) (Figure 6.6(c)) performs slightly better than with the Mangasarian function (17) (Figure 6.6(a)) and the Geman-McClure function (18) (Figure 6.6(b)).



(a) Mangasarian (b) Geman-McClure (c) Log-Det

Figure 6.6: Average SNR values vs. m/n for fixed $n = 1000$ and $s = 10$.

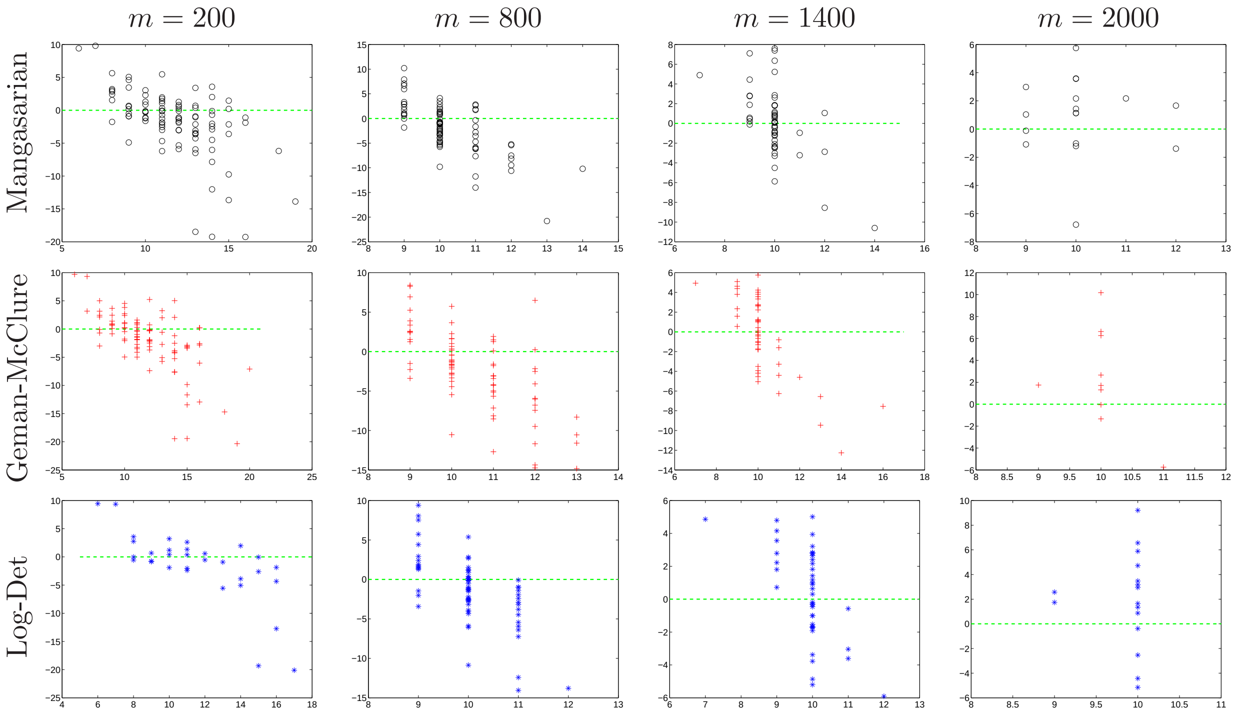
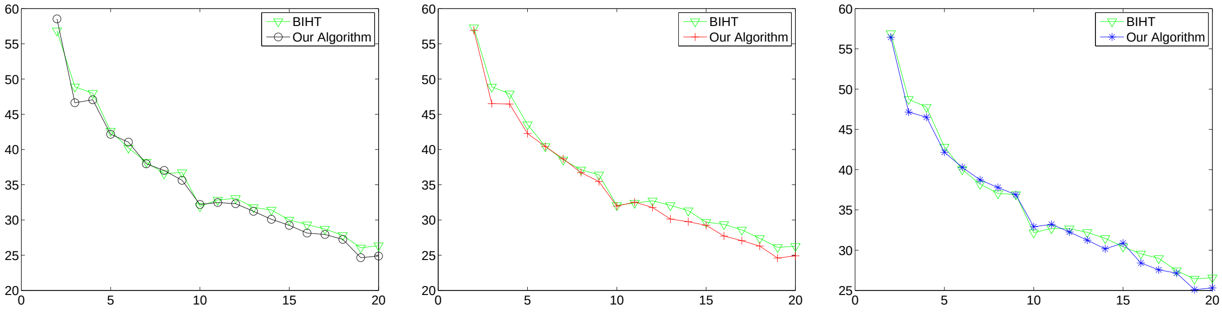


Figure 6.7: The difference of SNR values of estimates by Algorithm 3 and the BIHT vs. sparsity of reconstructed estimates by Algorithm 3. We fix $n = 1000$ and $s = 10$.

Detailed descriptions for valid trials for $m = 200$, 800, 1400, and 2000 are displayed in the columns (from left to right) of Figure 6.7, respectively. The plots in the rows (from the top to the bottom) of Figure 6.7 illustrate the results by Algorithm 3 with the Mangasarian function (17), the Geman-McClure function (18), and the Log-Det function (20), respectively. The horizontal axis of each plot represents the sparsity of the reconstructed signals by Algorithm 3 while the vertical axis represents the difference of the SNR values of the reconstructions between Algorithm 3 and the BIHT. Therefore, the marks (“$\circ$”, “$+$”, and “$\star$”) above the dashed horizontal lines indicate that Algorithm 3 performs better than the BIHT for the corresponding trials. Since all ideal signals in our simulations are 10-sparse, the marks whose horizontal axis are bigger than, exactly equal to, or smaller than 10 imply that the ℓ_0-norm of the reconstructions by Algorithm 3 are bigger than, exactly equal to, or smaller than 10, respectively. Thus, the ℓ_0-norm of a reconstruction over 10 indicates that the reconstruction is not a global minimizer of model (9), the ℓ_0-norm of a reconstruction being 10 indicates that the sparsity of the reconstruction is consistent with the one of the original test signal, the ℓ_0-norm of a reconstruction below 10 indicates that the reconstruction is potentially a global minimizer of model (9) and the original test signal is not a solution to model (9). We can conclude from Figure 6.7 that (i) the reconstructions by Algorithm 3 with sparsity higher (res. lower) than 10 usually have lower (res. higher) SNR values than that by the BIHT; (ii) Increasing m (number of measurements) tends to reduce the sparsity of the reconstructions. For example, average sparsity of the reconstructions for $m = 200$, 800, 1400, and 2000 are, respectively, 11.74, 10.26, 10, and 10.06 for Algorithm 3 with the Mangasarian function, 11.87, 10.53, 10.27, and 10 for Algorithm 3 with the Geman-McClure function, and 11.53, 10.05, 9.88, 9.88 for Algorithm 3 with the Log-Det function.

For the second configuration, the SNR values in decibels of the average reconstruction errors by both the BIHT and Algorithm 3 are compared in Figure 6.8 for varying sparsity of original signals. The plots demonstrate that our proposed algorithm performs better than the BIHT for sparsity s being 2 and 6 to 10. We emphasize again that unlike the BIHT the exact sparsity of the original signal is not required in advance by Algorithm 3. We remark that when $s = 1$ both the BIHT and Algorithm 3 find an exact solution to model (9). This phenomenon was also reported in [13]. Detailed descriptions for valid trials for $s = 2$, 8, 14, and 20 are displayed in the columns (from left to right) of Figure 6.9, respectively. The marks in each plot of Figure 6.9 have the same meaning as that in Figure 6.7. For fixed $m = 1000$ and $n = 1000$ we can draw conclusions from Figure 6.9 that (i) Algorithm 3 tends to produce an estimate whose sparsity is consistent with the ideal sparse signal; (ii) Algorithm 3 can give an estimate whose sparsity is smaller than that of the ideal sparse signal, in particular, when the sparsity of an original signal is relative large.



(a) Mangasarian

(b) Geman-McClure

(c) Log-Det

Figure 6.8: Average SNR values vs. sparsity of original signals for fixed $n = 1000$ and $m = 1000$.

For the third configuration, the SNR values in decibels of the average reconstruction errors by both the BIHT and Algorithm 3 are compared in Figure 6.10 for fixed $m = 1000$ and $s = 10$ and varying dimensions of original signals. The plots in Figure 6.10 show that the average SNR values for reconstructions by Algorithm 3 are lower than that by the BIHT in most cases. This is due to the fact that the BIHT explores an unattainable additional information on the sparsity of the

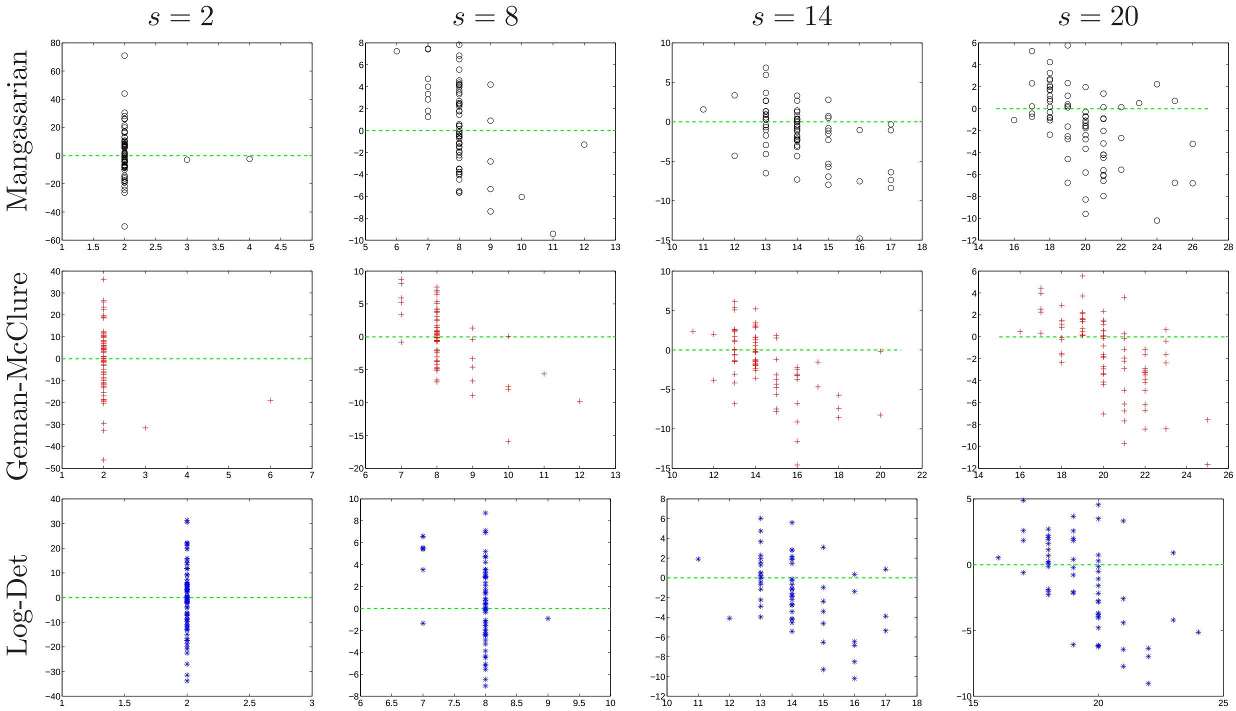
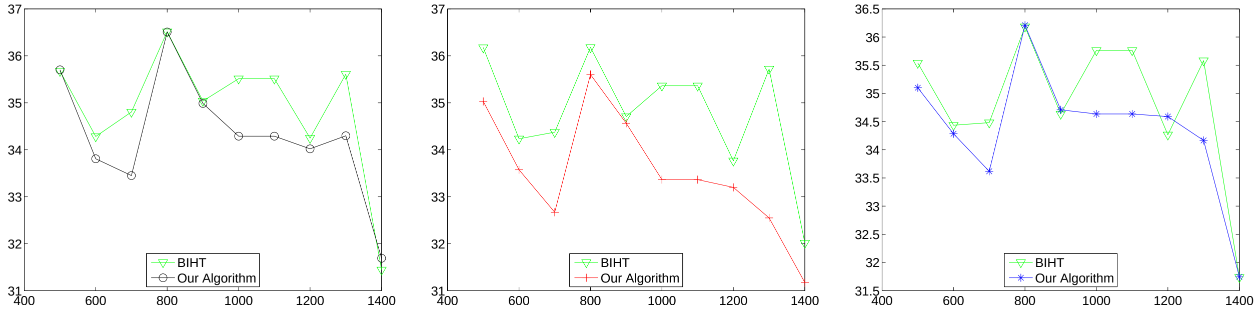


Figure 6.9: The difference of SNR values of estimates by Algorithm 3 and the BIHT vs. sparsity of reconstructed estimates by Algorithm 3. We fix $n = 1000$ and $m = 1000$.

original signal. Another reason which we can see from Figure 6.11 is that reconstructions with their sparsity larger than 10 by Algorithm 3 usually have lower SNR values than by the BIHT. The marks in each plot of Figure 6.11 have the same meaning as that in Figures 6.7 and 6.9. For fixed $m = 1000$ and $s = 10$ we can draw conclusions from Figure 6.11 that (i) Algorithm 3 can give an estimate whose sparsity is smaller than that of the ideal sparse signal and (ii) Algorithm 3 with the Log-Det function works better than that with the other two.



(a) Mangasarian (b) Geman-McClure (c) Log-Det

Figure 6.10: Average SNR values of estimates vs. the signal size n for fixed $m = 1000$ and $s = 10$.

7 Summary and Conclusion

In this paper we proposed a new model and algorithm for 1-bit compressive sensing. Unlike the state-of-the-art BIHT method, our model does not need to know the sparsity of the signal of

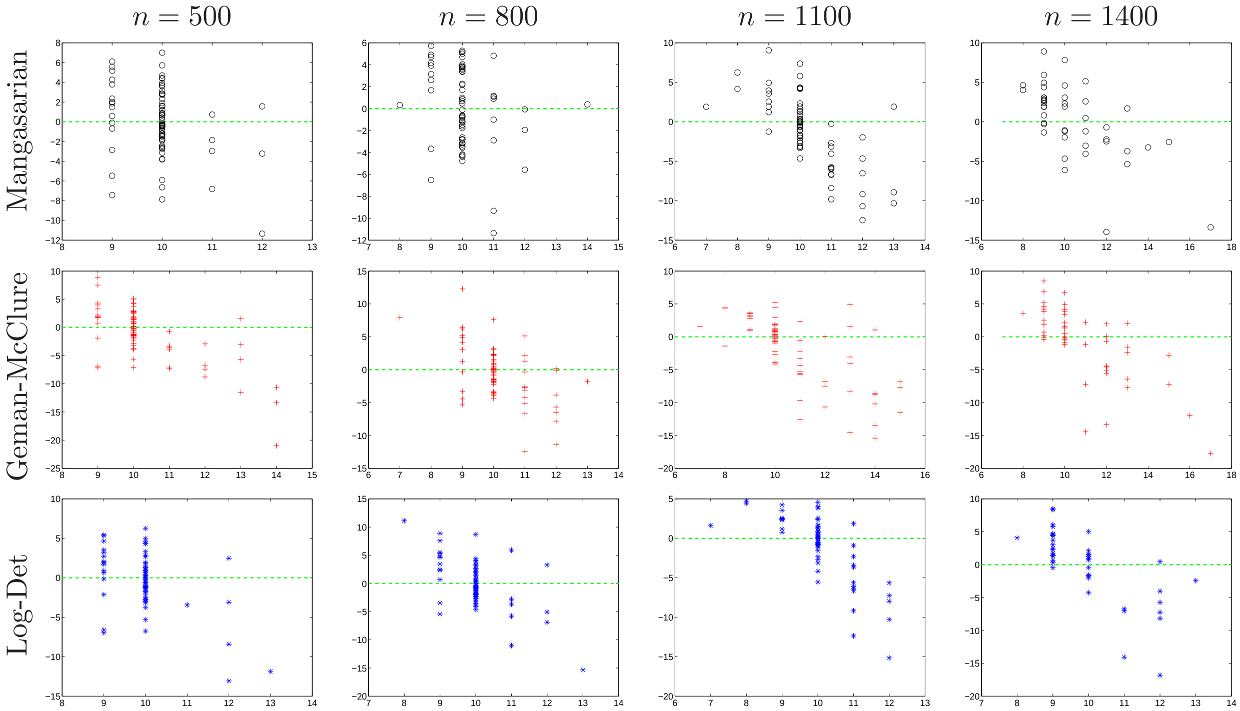


Figure 6.11: The difference of SNR values of estimates by Algorithm 3 and the BIHT vs. sparsity of reconstructed estimates by Algorithm 3. We fix $m = 1000$ and $s = 10$.

interest. We demonstrated the performance of our proposed algorithm for reconstruction from 1-bit measurements.

It would be of interest to study the convergence of Algorithm 3 with the Mangasarian function and the Geman-McClure function in the future. It would be highly needed to adaptively update all parameters in Algorithm 3 so that consistent reconstruction can be always achieved with improved accuracy of the solution.

Acknowledgements

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

References

[1] H. L. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, AMS Books in Mathematics, Springer, New York, 2011.

[2] T. Blumensath and M. E. Davies, *Iterative hard thresholding for compressed sensing*, Applied and Computational Harmonic Analysis, 27 (2009), pp. 265 – 274.

[3] P. T. Boufounos and R. G. Baraniuk, *1-bit compressive sensing*, Proceedings of Conference on Information Science and Systems (CISS), Princeton, NJ, 2008.

[4] E. Candes, J. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Communications on Pure and Applied Mathematics, 59 (2006), pp. 1207–1223.

[5] E. Candes and T. Tao, *Near optimal signal recovery from random projections: Universal encoding strategies?*, IEEE Transactions on Information Theory, 52 (2006), pp. 5406–5425.

[6] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 120–145.

[7] K. Davidson and S. Szarek, *Local operator theory, random matrices and Banach spaces*, in Handbook of the Geometry of Babach Spaces, vol. I, Amsterdam: North-Holland, 2001, pp. 317–366.

[8] M. Fazel, H. Hindi, and S. Boyd, *Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices*, vol. 3 of Proceedings of American Control Conference, 2003, pp. 2156 – 2162.

[9] D. Geman and G. Reynolds, *Nonlinear image recovery with half-quadratic regularization*, IEEE Transactions on Imae Processing, 4 (1995), pp. 932–946.

[10] E. Hale, W. Yin, and Y. Zhang, *Fixed-point continuation for ℓ_1 minimization: Methodology and convergence*, SIAM Journal on Optimization, 19 (2008), pp. 1107–1130.

[11] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, *Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors*, IEEE Transactions on Information Theory, accepted, (2012).

[12] S. Jokar and M. E. Pfetsch, *Exact and approximate sparse solutions of underdetermined linear equations*, SIAM Journal on Scientific Computing, 31 (2008), pp. 23–44.

[13] J. Laska, Z. Wen, W. Yin, and R. Baraniuk, *Trust, but verify: Fast and accurate signal recovery from 1-bitcompressive measurements*, IEEE Transactions on Signal Processing, 59 (2011), pp. 5289 – 5301.

[14] Q. Li, C. A. Micchelli, L. Shen, and Y. Xu, *A proximity algorithm accelerated by Gauss-Seidel iterations for* L1/TV *denoising models*, Inverse Problems, 28 (2012), p. 095003.

[15] O. L. Mangasarian, *Minimum-support solutions of polyhedral concave programs*, Optimization, 45 (1999), pp. 149 – 162.

[16] J.-J. Moreau, *Fonctions convexes duales et points proximaux dans un espace hilbertien*, C.R. Acad. Sci. Paris Sér. A Math., 255 (1962), pp. 1897–2899.

[17] ——, *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France, 93 (1965), pp. 273–299.

[18] M. Nikolova, *Local strong homogeneity of a regularized estimator*, SIAM Journal of Applied Mathematics, 61 (2000), pp. 633–658.

[19] Y. Plan and R. Vershynin, *One-bit compressed sensing by linear programming*, Communications on Pure and Applied Mathematics, accepted, (2012).

[20] L. Shen and B. W. Suter, *Bounds for eigenvalues of arrowhead matrices and their applications to hub matrices and wireless communications*, EURASIP Journal on Advances in Signal Processing, Article ID 379402, 12 pages, (2009).

[21] M. Yan, Y. Yang, and S. Osher, *Robust 1-bit compressive sensing using adaptive outlier pursuit*, IEEE Transactions on Signal Processing, 60 (2012), pp. 3868 – 3875.

[22] Y.-B. Zhao and D. Li, *Reweighted ℓ_1-minimization for sparse solutions to underdetermined linear system*, SIAM Journal on Optimization, 22 (2012), pp. 1065–1088.